Deep Learning Approaches for Water Quality Prediction in Aquaponics Systems: A Comparative Study of Recurrent and Feedforward Architectures

Gregorius Airlangga [1], **Oskar Ika Adi Nugroho** [2], **Lai Ferry Sugianto** [3]

[1] Department of information Systems, Atma Jaya Catholic University of Indonesia, Jakarta, Indonesia
[2] Department of Electrical Engineering, National Chung Cheng University, Chiayi, Taiwan
[3] Department of Business Administration, Fujen Catholic University, Taipei, Taiwan

ARTICLE INFORMATION

ABSTRACT
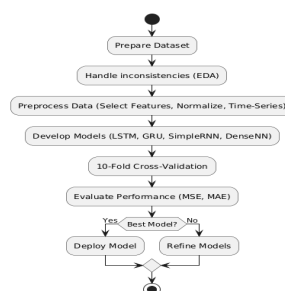
Accurate prediction of water quality parameters is critical for the effective management and sustainability of aquaponics systems. This study evaluates the performance of four deep learning architectures: Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Simple Recurrent Neural Network (SimpleRNN), and Dense Neural Network (DenseNN) for forecasting key water quality parameters, including temperature, turbidity, dissolved oxygen, pH, ammonia, and nitrate. A significant research gap is addressed by analyzing how these models perform on noisy and minimally preprocessed datasets, advancing prior studies that lack robust preprocessing techniques tailored for aquaponics systems. A ten-fold cross-validation framework was employed to rigorously assess the models, with Mean Squared Error (MSE) and Mean Absolute Error (MAE) as evaluation metrics. The results demonstrate that LSTM and GRU models outperform other architectures, achieving average validation losses of 0.0028 and 0.0028, respectively, and mean absolute errors of 0.0473 and 0.0478. These models effectively capture the temporal dependencies inherent in time-series data, making them highly suitable for the complex dynamics of aquaponics systems. Unlike previous studies, this research highlights the trade-offs between computational efficiency and predictive accuracy in these models. In contrast, the SimpleRNN model exhibited higher error rates due to its inability to model long-term dependencies, while the DenseNN model, lacking temporal processing mechanisms, showed the lowest performance with an average validation loss of 0.0075 and MAE of 0.0797. This study underscores the importance of selecting appropriate model architectures for time-series forecasting tasks and provides a foundation for deploying predictive systems to optimize aquaponics operations. Future work includes exploring hybrid models with attention mechanisms and real-time data integration for enhanced operational efficiency.

Document Citation:

1. INTRODUCTION

Aquaponics, a symbiotic system combining aquaculture and hydroponics, has emerged as a sustainable agricultural approach that offers efficient resource utilization and environmental benefits [1]–[3]. However, the effective management of aquaponics systems depends heavily on maintaining optimal water quality parameters such as pH, ammonia, nitrate, and dissolved oxygen [4]–[6]. These parameters are critical for ensuring the health of aquatic organisms and the overall productivity of the system [7]. Deviations from optimal conditions can lead to adverse outcomes, including fish mortality, reduced plant growth, and system imbalances [8]. Despite the increasing adoption of aquaponics systems globally, maintaining the delicate balance of these parameters remains a significant challenge due to the dynamic and interdependent nature of the system [5],[9],[10]. The complexity of monitoring and predicting water quality in aquaponics systems is further compounded by external factors such as environmental variations, sensor inaccuracies, and system-specific characteristics [11]–[13]. Research into real-time predictive models for water quality has gained momentum, but existing studies often lack comprehensive evaluations across multiple architectures or fail to address the variability and noise inherent in aquaponics datasets. Traditional methods for monitoring rely on manual observation or rule-based systems, which are often inadequate for real-time or predictive decision-making [14]. As a result, researchers have turned to advanced machine learning and deep learning techniques to address these challenges. Deep learning models, in particular, have shown immense potential in capturing non-linear relationships and temporal dependencies within complex datasets, making them suitable for forecasting water quality in aquaponics systems [14]–[16].

Several studies have highlighted the applicability of machine learning in environmental monitoring and time-series forecasting. For instance, [17] demonstrated the use of LSTM networks for predicting water quality in fish ponds, achieving notable accuracy in forecasting dissolved oxygen levels. Similarly, [18] utilized GRU models to analyze the impact of temperature and ammonia on aquaculture systems, showcasing the model's ability to handle sequential dependencies effectively. However, these studies often focus on isolated models and overlook comparative performance under rigorous preprocessing and evaluation frameworks. Another study by [19][20] compared traditional statistical methods with neural network-based approaches for pH level prediction in hydroponics, concluding that deep learning models significantly outperformed conventional methods in terms of precision and adaptability. Despite these advancements, there remains a gap in the literature regarding the comparative performance of various deep learning architectures for aquaponics water quality prediction, particularly when dealing with noisy and minimally processed datasets [21]–[23]. Moreover, few studies have explored the integration of preprocessing pipelines tailored specifically for aquaponics datasets to enhance model robustness and generalizability [15],[24][25]. Addressing these gaps is crucial for developing practical and scalable solutions that can be deployed in real-world aquaponics systems.

This study aims to bridge these gaps by systematically evaluating the performance of four state-of-the-art deep learning models: Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Simple Recurrent Neural Network (SimpleRNN), and Dense Neural Network (DenseNN) for forecasting key water quality parameters in aquaponics systems. A rigorous preprocessing pipeline is developed to handle noise, missing values, and feature scaling, ensuring that the models are trained on clean and reliable data. The models are evaluated using a ten-fold cross-validation approach to ensure comprehensive performance assessment and generalizability. The primary contributions of this research are threefold. First, it provides a comparative analysis of advanced deep learning models for forecasting water quality in aquaponics systems under realistic conditions. Second, it introduces a robust and reproducible preprocessing framework tailored for noisy aquaponics datasets, addressing challenges related to data inconsistencies and feature alignment. Third, it delivers actionable insights into the trade-offs among different model architectures, equipping practitioners with the knowledge to make informed decisions when implementing predictive models in aquaponics systems. The remainder of this paper is organized as follows. Section 2 details the materials and methods, including the preprocessing pipeline and model architectures. Section 3 presents the experimental results, followed by a discussion of the findings. Finally, Section 4 concludes the study and outlines potential directions for future research.

2. METHODS

This section elaborates on the comprehensive methodology utilized in this study, encompassing dataset preparation, data preprocessing, model development, cross-validation, and performance evaluation as presented in Figure 1. The methodology emphasizes reproducibility and robustness to address the challenges posed by noisy datasets.
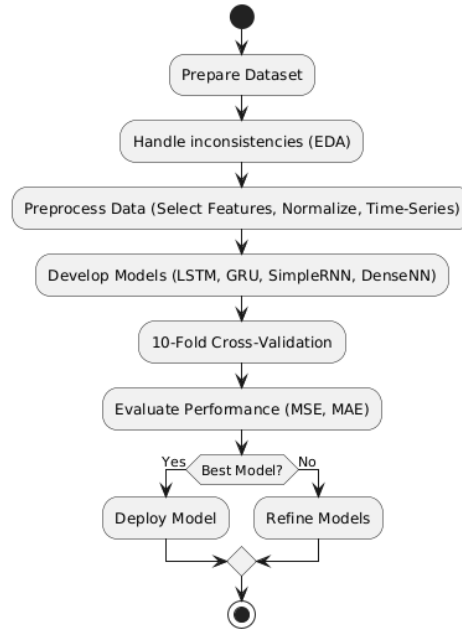
Figure 1. Research methodology

2.1. Dataset Preparation

The dataset used in this research was sourced from [26], containing time-series data collected through sensors installed in aquaponics fish ponds. The dataset includes key water quality parameters such as temperature $((T))$, turbidity $((Tu))$, dissolved oxygen $((DO))$, pH $((pH))$, ammonia $((NH_3))$, and nitrate $((NO_3))$. Formally, the dataset can be expressed as a collection of pairs:

$$D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\},$$

where $(x_i \in R^d)$ represents the (d)-dimensional feature vector for observation (i), and $(y_i \in R)$ represents the target variable for prediction. The dataset, composed of multiple CSV files, was unified into a single structured file to facilitate analysis. Exploratory Data Analysis (EDA) was conducted to inspect the dataset for inconsistencies such as missing values, outliers, and feature correlations. Statistical summaries and visualization techniques were employed to understand the distribution and interdependencies of the variables.

2.2. Data Preprocessing

Data preprocessing was critical due to the noisy and minimally cleaned nature of the dataset. The following steps were meticulously carried out to prepare the data for model training. Firstly, feature selection and alignment, in this stage, we create a subset of features relevant to aquaponics water quality monitoring was selected based on domain knowledge. These features were aligned to ensure consistency between expected variables and the actual dataset. Let the selected features be represented as (1).

$$\mathcal{F}_{selected} = \{T, Tu, DO, pH, NH_3, NO_3\} \tag{1}$$

An algorithm was implemented to match the selected features against the dataset's actual columns, correcting discrepancies where necessary. Second, handling missing and invalid data, in this state, rows with missing or invalid values were identified and handled. Specifically, infinite values $((\pm\infty))$ were replaced with NaN and subsequently removed by following rule as presented in (2).

$$D_{clean} = \{(x_i, y_i) \mid x_i \text{ contains no NaN or } \pm \infty\} \tag{2}$$

Third is Normalization, in this stage, feature scaling was performed to normalize the dataset into the range $([0,1])$ using the Min-Max Scaling formula as presented in (3).

$$x'_{i,j} = \frac{x_{i,j} - min(x_j)}{max(x_j) - min(x_j)} \tag{3}$$

where $(x_{i,j})$ is the value of feature (j) in observation (i), and $(\min(x_j))$, $(\max(x_j))$ are the minimum and maximum values of feature (j) across all observations. Lastly we use Time-Series Sequence Creation method in order to capture temporal dependencies, sequences of historical observations were constructed using a sliding window approach as presented in (4).

$$X_t = \{x_{t-w+1}, x_{t-w+2}, \dots, x_t\}, \quad y_t = x_{t+1} \tag{4}$$

where (w) is the window size, (X_t) is the sequence of (w) historical data points, and (y_t) is the target value for time (t). This transformation converts the dataset into sequences suitable for time-series forecasting models.

2.3. Model Development

Four deep learning architectures were developed to predict water quality parameters: Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Simple Recurrent Neural Network (SimpleRNN), and Dense Neural Network (DenseNN). Each model architecture was tailored to leverage unique characteristics of the data. The LSTM model utilizes gating mechanisms to retain long-term dependencies in sequential data. The update equations for the LSTM gates are presented in (5) to (10).

$$f_t = \sigma\big(W_f \cdot [h_{t-1}, x_t] + b_f\big), \tag{5}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \tag{6}$$

$$\widetilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \tag{7}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \widetilde{c}_t, \tag{8}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \tag{9}$$

$$h_t = o_t \odot \tanh(c_t), \tag{10}$$

where (f_t, i_t, o_t) represent the forget, input, and output gates, respectively, and (c_t, h_t) represent the cell and hidden states. The GRU simplifies these operations by directly updating the hidden state (h_t). The SimpleRNN model updates its hidden state iteratively as (11).

$$h_t = \sigma(W \cdot h_{t-1} + U \cdot x_t + b) \tag{11}$$

The DenseNN, a feedforward network, is expressed as $y = \sigma(W_2 \cdot \sigma(W_1 \cdot x + b_1) + b_2)$, where (W_1, W_2) and (b_1, b_2) represent weights and biases of the layers, and (σ) is the activation function.

2.4. Cross-Validation

A ten-fold cross-validation strategy was employed to assess model performance. The dataset was partitioned into ten subsets. For each fold (k), one subset served as the validation set $((D_k^{val}))$, and the remaining nine subsets formed the training set $((D_k^{train}))$. Model evaluation metrics for each fold were calculated as (12) and (13).

$$MSE_k = \frac{1}{|D_k^{val}|} \sum_{(x_i, y_i) \in D_k^{val}} (y_i - \widehat{y_i})^2, \tag{12}$$

$$MAE_k = \frac{1}{|D_k^{val}|} \sum_{(x_i, y_i) \in D_k^{val}} |y_i - \widehat{y_i}|, \tag{13}$$

where $(\widehat{y_i})$ is the predicted value. After that, in order to do performance evaluation, the models were evaluated using Mean Squared Error (MSE) and Mean Absolute Error (MAE), averaged across folds: $MSE_{avg} = \frac{1}{k} \sum_{k=1}^{10} MSE_k$ and $MAE_{avg} = \frac{1}{k} \sum_{k=1}^{10} MAE_k$.

2.5. Implementation Framework

The entire workflow was implemented in Python, utilizing TensorFlow for deep learning, Scikit-learn for preprocessing and evaluation, and Matplotlib for visualization. GPU-enabled environments were employed to accelerate computations. This comprehensive methodology provides a robust framework for exploring the

effectiveness of advanced deep learning models in forecasting water quality parameters in aquaponics systems, addressing the unique challenges posed by noisy and minimally preprocessed datasets.

3.   RESULT AND DISCUSSION

The results of this study provide an in-depth evaluation of four advanced deep learning models: LSTM, GRU, SimpleRNN, and DenseNN for forecasting water quality parameters in aquaponics systems as presented in the Table 1 to Table 5. These models were assessed using a rigorous ten-fold cross-validation approach, with Mean Squared Error (MSE) and Mean Absolute Error (MAE) serving as the primary metrics. This section elaborates on the performance of each model, discusses observed trends, and interprets the implications of these findings in detail. The LSTM model demonstrated robust performance across all folds, with an average validation loss of 0.0028 and an average MAE of 0.0473. The model achieved its best performance in Fold 3, where the MAE was as low as 0.0101. This exceptional accuracy indicates the model's capability to capture and generalize patterns in the data under favorable conditions. However, there were folds where performance varied slightly, with the MAE reaching 0.0723 in Fold 9. Such variability reflects the sensitivity of LSTM to data distribution and partitioning, particularly when dealing with noisy and imbalanced subsets. The ability of LSTM to retain long-term dependencies through its gating mechanisms contributes significantly to its performance in modeling the temporal dynamics of water quality parameters. This makes it particularly effective for datasets characterized by complex sequential relationships.

The GRU model achieved results comparable to LSTM, with an average validation loss of 0.0028 and an average MAE of 0.0478. Similar to LSTM, GRU displayed variability in performance across folds, with its best MAE recorded at 0.0198 in Fold 7 and its worst at 0.0748 in Fold 3. The variability suggests that while GRU is efficient and computationally lighter than LSTM due to its simpler architecture, it may struggle slightly more with certain subsets of the data. The GRU model's strength lies in its ability to maintain performance while reducing computational overhead, making it an attractive choice for scenarios where resource constraints exist. Despite these advantages, the results indicate that GRU's simplified architecture may occasionally fall short in capturing finer temporal details present in highly dynamic datasets.

The SimpleRNN model, with an average validation loss of 0.0037 and an average MAE of 0.0553, performed less effectively compared to LSTM and GRU. The model exhibited pronounced variability, with its best MAE recorded at 0.0363 in Fold 2 and its worst at 0.0801 in Fold 6. This wide range of performance highlights the limitations of SimpleRNN, which lacks the sophisticated gating mechanisms of LSTM and GRU. As a result, it is more prone to issues like vanishing gradients, particularly in tasks requiring the modeling of long-term dependencies. The relatively higher errors suggest that SimpleRNN struggled to capture the intricate temporal patterns in the aquaponics dataset. This makes it less suitable for applications involving time-series forecasting, where temporal dependencies are critical for accurate predictions. The DenseNN model, which uses a feedforward architecture, exhibited the lowest overall performance, with an average validation loss of 0.0075 and an average MAE of 0.0797. Although the model showed reasonable accuracy in some folds, such as Fold 1 where the MAE was 0.0382, it struggled significantly in others, with the MAE reaching 0.1008 in Fold 5. The lack of recurrent connections in DenseNN renders it incapable of modeling sequential dependencies, which are essential for forecasting tasks involving time-series data. Instead, DenseNN relies solely on the static relationships among features, limiting its effectiveness in scenarios where temporal dynamics play a crucial role. The higher error rates observed for DenseNN underscore the importance of using architectures specifically designed for time-series data when addressing problems such as water quality prediction.

A comparative analysis of the models reveals that LSTM and GRU outperformed SimpleRNN and DenseNN in both accuracy and generalizability. The results demonstrate the critical importance of recurrent architectures in capturing the temporal characteristics of the dataset. The strong performance of LSTM, coupled with its slightly lower variability across folds compared to GRU, suggests that it is the most robust model for this application. However, GRU remains a competitive alternative, particularly in scenarios where computational efficiency is a priority. The findings highlight the need for models that are well-aligned with the nature of the data being analyzed. The stark contrast between the performance of DenseNN and the recurrent models reinforces the importance of leveraging temporal dependencies in time-series forecasting. Additionally, the variability observed across folds for all models underscores the significance of robust data preprocessing and augmentation techniques to mitigate the impact of noise and ensure consistent performance. Future studies could explore advanced techniques such as domain-specific data augmentation, denoising, and feature engineering to further enhance model robustness and accuracy.

The superior performance of LSTM and GRU models underscores their potential for real-world deployment in aquaponics systems. These models can provide accurate and timely predictions of water quality parameters, enabling proactive interventions to maintain optimal conditions. The scalability and adaptability

of these models make them suitable for integration into automated monitoring systems, supporting sustainable aquaponics practices. In conclusion, the results of this study emphasize the effectiveness of LSTM and GRU models for forecasting water quality in aquaponics systems. The ability of these models to handle complex temporal dynamics positions them as valuable tools for predictive analytics in environmental monitoring. This research provides a strong foundation for future investigations aimed at refining these methodologies and exploring their applicability across diverse domains. The insights gained from this study contribute to advancing the state of the art in time-series forecasting and its applications in sustainable agriculture.

Table 1. Results for LSTM

Fold	Validation Loss	MAE
1	0.0039	0.0588
2	0.0028	0.0511
3	0.0002	0.0101
4	0.0024	0.0470
5	0.0012	0.0339
6	0.0049	0.0652
7	0.0025	0.0477
8	0.0021	0.0438
9	0.0060	0.0723
10	0.0021	0.0435

Table 2. Results for GRU

Fold	Validation Loss	MAE
1	0.0019	0.0428
2	0.0027	0.0509
3	0.0064	0.0748
4	0.0018	0.0418
5	0.0039	0.0610
6	0.0024	0.0475
7	0.0006	0.0198
8	0.0027	0.0505
9	0.0045	0.0621
10	0.0008	0.0267

Table 3. Results for SimpleRNN

Fold	Validation Loss	MAE
1	0.0042	0.0601
2	0.0014	0.0363
3	0.0037	0.0579
4	0.0042	0.0602
5	0.0034	0.0559
6	0.0073	0.0801
7	0.0045	0.0621
8	0.0016	0.0393
9	0.0046	0.0633
10	0.0015	0.0376

Table 4. Results for DenseNN

Fold	Validation Loss	MAE
1	0.0015	0.0382
2	0.0104	0.0970
3	0.0078	0.0830
4	0.0083	0.0855
5	0.0111	0.1008
6	0.0066	0.0769
7	0.0055	0.0696
8	0.0102	0.0959
9	0.0036	0.0566
10	0.0097	0.0935

Table 5. Performance Results

Model	Average Validation Loss	Average MAE
LSTM	0.0028	0.0473
GRU	0.0028	0.0478
RNN	0.0037	0.0553
DenseNet	0.0075	0.0797

4. CONCLUSIONS

This study comprehensively evaluated the performance of four deep learning architectures: LSTM, GRU, SimpleRNN, and DenseNN for forecasting water quality parameters in aquaponics systems. The models were rigorously tested using a ten-fold cross-validation framework, with Mean Squared Error (MSE) and Mean Absolute Error (MAE) as the primary metrics. The results demonstrate that recurrent architectures, particularly LSTM and GRU, outperformed other models, underscoring their effectiveness in capturing temporal dependencies inherent in time-series data. The LSTM model exhibited the most consistent performance, achieving an average validation loss of 0.0028 and an average MAE of 0.0473. Its robust architecture and ability to retain long-term dependencies make it highly suitable for complex and noisy datasets like those used in this study. Similarly, the GRU model, with an average validation loss of 0.0028 and an average MAE of 0.0478, provided comparable results, offering a computationally efficient alternative to LSTM. The SimpleRNN model, while simpler, showed limitations in handling long-term dependencies, resulting in higher error rates with an average validation loss of 0.0037 and an average MAE of 0.0553. The DenseNN model, which lacks a mechanism to process temporal patterns, demonstrated the lowest accuracy, with an average validation loss of 0.0075 and an average MAE of 0.0797, highlighting the importance of recurrent structures for time-series forecasting tasks.

The findings of this study emphasize the critical role of model architecture in addressing the unique challenges of forecasting water quality parameters in aquaponics systems. Recurrent models, particularly LSTM and GRU, are shown to be effective in capturing the complex dynamics of such systems, providing actionable insights for maintaining optimal water conditions. Furthermore, the results underscore the importance of robust preprocessing techniques to address data noise and ensure reliable predictions. This research provides a foundation for future studies aimed at optimizing predictive models for aquaponics systems. Future work could explore advanced ensemble techniques, hybrid architectures, and domain-specific data augmentation methods to further improve model accuracy and generalizability. Additionally, integrating real-time sensor data with predictive systems could enhance the operational efficiency of aquaponics systems, supporting sustainable agricultural practices. Through this study, the potential of deep learning in transforming aquaponics monitoring and management has been demonstrated, paving the way for more sophisticated and impactful applications in the field.

REFERENCES

[1] L. A. Ibrahim, H. Shaghaleh, G. M. El-Kassar, M. Abu-Hashim, E. A. Elsadek, and Y. Alhaj Hamoud, "Aquaponics: a sustainable path to food sovereignty and enhanced water use efficiency," *Water*, vol. 15, no. 24, p. 4310, 2023, https://doi.org/10.3390/w15244310.

[2] B. Yep and Y. Zheng, "Aquaponic trends and challenges--A review," *J. Clean. Prod.*, vol. 228, pp. 1586–1599, 2019, https://doi.org/10.1016/j.jclepro.2019.04.290.

[3] M. Schoor, A. P. Arenas-Salazar, I. Torres-Pacheco, R. G. Guevara-González, and E. Rico-Garcia, "A review of sustainable pillars and their fulfillment in Agriculture, aquaculture, and Aquaponic Production," *Sustainability*, vol. 15, no. 9, p. 7638, 2023, https://doi.org/10.3390/su15097638.

[4] A. R. Yanes, P. Martinez, and R. Ahmad, "Towards automated aquaponics: A review on monitoring, IoT, and smart systems," *J. Clean. Prod.*, vol. 263, p. 121571, 2020, https://doi.org/10.1016/j.jclepro.2020.121571.

[5] P. Debroy, P. Majumder, and L. Seban, "A simulation based water quality parameter control of aquaponic system employing model predictive control strategy incorporation with optimization technique," *Environ. Prog. & Sustain. Energy*, p. e14530, 2024, https://doi.org/10.1002/ep.14530.

[6] P. Chandramenon, A. Aggoun, and F. Tchuenbou-Magaia, "Smart approaches to Aquaponics 4.0 with focus on water quality- Comprehensive review," *Comput. Electron. Agric.*, vol. 225, p. 109256, 2024, https://doi.org/10.1016/j.compag.2024.109256.

[7] A. B. Dauda, "Biofloc technology: a review on the microbial interactions, operational parameters and implications to disease and health management of cultured aquatic animals," *Rev. Aquac.*, vol. 12, no. 2, pp. 1193–1210, 2020, https://doi.org/10.1111/raq.12379.

[8] M. J. Islam, A. Kunzmann, and M. J. Slater, "Responses of aquaculture fish to climate change-induced extreme temperatures: A review," *J. World Aquac. Soc.*, vol. 53, no. 2, pp. 314–366, 2022, https://doi.org/10.1111/jwas.12853.

[9] A. K. Verma, M. H. Chandrakant, V. C. John, R. M. Peter, and I. E. John, "Aquaponics as an integrated agri-aquaculture system (IAAS): Emerging trends and future prospects," *Technol. Forecast. Soc. Change*, vol. 194, p. 122709, 2023, https://doi.org/10.1016/j.techfore.2023.122709.

[10] T. Malmir. *System Dynamics Modeling of the Food-Water-Energy Nexus in Urban Areas, Focusing on Community Gardens.* Concordia University Montreal, Quebec, Canada, 2023, https://spectrum.library.concordia.ca/id/eprint/992768/.

[11] Z. Schmautz. *Characterization of nitrogen dynamics in an aquaponic system*. (Doctoral dissertation, ETH Zurich), 2020, https://www.research-collection.ethz.ch/handle/20.500.11850/464851.

[12] A. Kobelski, P. Nestler, M. Mauerer, T. Rocksch, U. Schmidt, and S. Streif, "An Algorithm for Nutrient Mixing Optimization in Aquaponics," *Appl. Sci.*, vol. 14, no. 18, p. 8140, 2024, https://doi.org/10.3390/app14188140.

[13] K. Kazimierczuk, S. E. Barrows, M. V Olarte, and N. P. Qafoku, "Decarbonization of Agriculture: The Greenhouse Gas Impacts and Economics of Existing and Emerging Climate-Smart Practices," *ACS Eng. Au*, vol. 3, no. 6, pp. 426–442, 2023, https://doi.org/10.1021/acsengineeringau.3c00031.

[14] J. Gladju, B. S. Kamalam, and A. Kanagaraj, "Applications of data mining and machine learning framework in aquaculture and fisheries: A review," *Smart Agric. Technol.*, vol. 2, p. 100061, 2022, https://doi.org/10.1016/j.atech.2022.100061.

[15] A. Khandakar *et al.*, "Smart aquaponics: An innovative machine learning framework for fish farming optimization," *Comput. Electr. Eng.*, vol. 119, p. 109590, 2024, https://doi.org/10.1016/j.compeleceng.2024.109590.

[16] A. E. Alprol, A. T. Mansour, M. E. E.-D. Ibrahim, and M. Ashour, "Artificial Intelligence Technologies Revolutionizing Wastewater Treatment: Current Trends and Future Prospective," *Water*, vol. 16, no. 2, p. 314, 2024, https://doi.org/10.3390/w16020314.

[17] J. Liu, C. Zhang, D. An, and Y. Wei, "Development and application of an innovative dissolved oxygen prediction fusion model," *Comput. Electron. Agric.*, vol. 227, p. 109496, 2024, https://doi.org/10.1016/j.compag.2024.109496.

[18] W. Li, H. Wu, N. Zhu, Y. Jiang, J. Tan, and Y. Guo, "Prediction of dissolved oxygen in a fishery pond based on gated recurrent unit (GRU)," *Inf. Process. Agric.*, vol. 8, no. 1, pp. 185–193, 2021, https://doi.org/10.1016/j.inpa.2020.02.002.

[19] M. A. Jabed and M. A. A. Murad, "Crop Yield Prediction in Agriculture: A Comprehensive Review of Machine Learning and Deep Learning Approaches, with Insights for Future Research and Sustainability," *Heliyon*, 2024, https://doi.org/10.1016/j.heliyon.2024.e40836.

[20] T. Van Klompenburg, A. Kassahun, and C. Catal, "Crop yield prediction using machine learning: A systematic literature review," *Comput. Electron. Agric.*, vol. 177, p. 105709, 2020, https://doi.org/10.1016/j.compag.2020.105709.

[21] A. Metin, A. Kasif, and C. Catal, "Temporal fusion transformer-based prediction in aquaponics," *J. Supercomput.*, vol. 79, no. 17, pp. 19934–19958, 2023, https://doi.org/10.1007/s11227-023-05389-8.

[22] J. Liu and W. Jiang, "Optimization of Aquaponics System Efficiency based on Artificial Intelligence Approach," in *2024 IEEE 3rd International Conference on Computing and Machine Intelligence (ICMI)*, pp. 1–6, 2024, https://doi.org/10.1109/ICMI60790.2024.10586162.

[23] S. B. Dhal *et al.*, "A machine-learning-based IoT system for optimizing nutrient supply in commercial aquaponic operations," *Sensors*, vol. 22, no. 9, p. 3510, 2022, https://doi.org/10.3390/s22093510.

[24] S. Kanwal *et al.*, "An Optimal Internet of Things-Driven Intelligent Decision-Making System for Real-Time Fishpond Water Quality Monitoring and Species Survival," *Sensors*, vol. 24, no. 23, p. 7842, 2024, https://doi.org/10.3390/s24237842.

[25] M. A. Rahu, A. F. Chandio, K. Aurangzeb, S. Karim, M. Alhussein and M. S. Anwar, "Toward Design of Internet of Things and Machine Learning-Enabled Frameworks for Analysis and Prediction of Water Quality," in *IEEE Access*, vol. 11, pp. 101055-101086, 2023,, https://doi.org/10.1109/ACCESS.2023.3315649.

[26] S. Fetouh, "Cleaned Aquaponics Pond Dataset." Kaggle, 2024, https://www.kaggle.com/datasets/samahfetouh/cleaned-aquaponics-pond-dataset.

AUTHOR BIOGRAPHY

Gregorius Airlangga is Received the B.S. degree in information system from the Yos Sudarso Higher School of Computer Science, Purwokerto, Indonesia, in 2014, and the M.Eng. degree in informatics from Atma Jaya Yogyakarta University, Yogyakarta, Indonesia, in 2016. He got Ph.D. degree with the Department of Electrical Engineering, National Chung Cheng University, Taiwan. He is also an Assistant Professor with the Department of Information System, Atma Jaya Catholic University of Indonesia, Jakarta, Indonesia. His research interests include UAV, data science, artificial intelligence and software engineering include path planning, machine learning, natural language processing, deep learning, software requirements, software design pattern and software architecture.